



Denmark, T., Fish, J., Jansari, A., Tailor, J., Ashkan, K. and Morris, R. (2019) Using Virtual Reality to investigate multitasking ability in individuals with frontal lobe lesions. *Neuropsychological Rehabilitation*, 29(5), pp. 767-788.

There may be differences between this version and the published version. You are advised to consult the publisher's version if you wish to cite from it.

<http://eprints.gla.ac.uk/203125/>

Deposited on: 28 November 2019

Enlighten – Research publications by members of the University of Glasgow
<http://eprints.gla.ac.uk>

Title: Using Virtual Reality to investigate multitasking ability in individuals with frontal lobe lesions

Authors. Tanya Denmark¹, Jessica Fish^{1, 2}, Ashok Jansari³, Jignesh Tailor⁴ Keyoumars Ashkan^{1,4*} and Robin Morris^{1,5*}

Affiliations.

¹Institute of Psychiatry, Psychology and Neuroscience, Kings College London, London, United Kingdom

²The Oliver Zangwill Centre for Neuropsychological Rehabilitation, Cambridgeshire Community Services NHS Trust, Ely, UK.

³Department of Psychology, Goldsmiths University of London, United Kingdom

⁴Department of Neurosurgery, Kings College Hospital, London, United Kingdom

⁵ Department of Neuropsychology, Kings College Hospital, London, United Kingdom

Short Title: Virtual Reality multitasking following frontal lobe lesions

Word count (excluding title page, abstract, references, figures and tables): TBC

*** Joint senior authors**

Corresponding author:

Correspondence should be addressed to Professor Robin Morris, PO BOX 78, Institute of Psychiatry, Psychology and Neuroscience, De Crespigny Park, London, SE5 8AF, E-mail: Robin.Morris@kcl.ac.uk.

ABSTRACT

Individuals with lesions in the prefrontal cortex often show impairments with the organisation of their behaviour in everyday life. These difficulties can be hard to detect using structured formal tests. The objective of this study was to use Virtual Reality (VR) to explore the multitasking performance of individuals with focal frontal lobe lesions, specifically using the Jansari assessment of Executive Functions (JEF[®] Jansari et al., 2014). Nineteen individuals with frontal lobe lesions were compared with 19 matched controls on the test and a group of commonly used clinical measures of neuropsychological functioning, as well as questionnaire measures of everyday activity, anxiety and depression. There was a significant difference between groups on the overall JEF[®] score and on five of the eight individual constructs, namely the planning, creative thinking, adaptive thinking, event-based Prospective Memory (PM) and time-based PM constructs. There were no differences between groups on the non-VR EF individual measures apart from on one EF control measure, Trail Making A. These results demonstrate the potential clinical utility of the JEF[®] and highlight the value of ecologically valid VR measures in detecting impairments in EF in individuals with frontal lobe lesions.

Keywords: Executive function; Prefrontal cortex; Virtual Reality; Ecologically valid; Neuropsychology.

INTRODUCTION

The term executive functioning (EF) refers to a set of cognitive abilities such as planning, initiation, goal management, prospective memory and self-monitoring, which can be flexibly used when individuals are faced with the multiple goals, sub-tasks and changing priorities commonly encountered in everyday life (Shallice, Burgess & Robertson, 1996). Many researchers have shown that the prefrontal cortex (PFC) significantly contributes to executive processes (e.g. Baddeley, 1986; Stuss and Benson, 1986; Elliott, 2003) and individuals who present with cognitive and behavioural impairment following damage to the PFC frequently present with a dysexecutive syndrome (Funahashi, 2001). Allied to EF is prospective memory (PM), remembering to perform an intended action in the future, a common element of many executive tasks (Ellis, 1996; Ellis & Freeman, 2008) and also supported by the PFC (Shallice and Burgess, 1991; Okuda, 1998; Neulinger, Oram, Tinson, O’Gorman & Shtum, 2016).

There are numerous neuropsychological procedures for measuring EF, including well-used measures such as the Wisconsin Card Sorting Test (WCST; e.g. Heaton, 1981; Nyhus & Barcelo, 2009) and the Stroop Test (e.g. Delis, Kaplan & Kramer, 2001) among many more. While such procedures are frequently used they often fail to detect EF impairment, particularly in individuals with PFC damage (Shallice, 1982; Anderson, Bigler & Blatter, 1995). The lack of sensitivity presents a problem for neuropsychological assessment and formulation and is likely to be due to the tests eliciting cognitive activity that is too constrained to reflect the type of EF difficulties associated with everyday activities (Eslinger & Damasio, 1985; Shallice & Burgess, 1991; Burgess et al., 1998; 2006). This so-called ‘frontal paradox’ (Shallice & Burgess, 1991) has led to efforts being made to develop new assessment measures that have greater ‘ecological validity’. A specific example of this is the Multiple Errands Test (MET) developed by Shallice and Burgess (1991) in a landmark study; they designed a shopping task, which requires individuals to undertake a series of errands, for example, buy specified items in a pedestrian precinct. More complex tasks were also included, such as obtaining the necessary items to send a postcard and certain fact-finding errands and specific rules to follow. Shallice and Burgess (1991) demonstrated that three individuals with frontal lobe injuries had impaired performance on the MET, despite relatively normal performance on other EF tests. Such findings have been replicated in other studies, showing the tendency of

individuals with PFC damage to have specific difficulties when applying efficient strategies in multi-tasking situations, but measured using simulation neuropsychological procedures (Goldstein, Bernard, Fenwick, Burgess, & McNeil, 1993; Crepeau, Belleville, & Duchesne, 1996; Bisiacchi, Sgaramella, & Farinello, 1998; Manly, Hawkins, Evans, Woldt & Robertson, 2002; Hsu, Zanto, Anguera, Lin & Gazzaley, 2015). Additionally, there are standardised EF procedures designed to mimic everyday EF activity, such as the Behavioural Assessment of the Dysexecutive Syndrome (BADS) test battery (Wilson, Alderman, Burgess, Emslie & Evans, 1996).

The ‘ecological’ approaches have tended to use either real world activity, which is time consuming, or ‘paper and pencil’ methodology to measure EF. With the advent of more powerful and flexible computing technology, however, there is now a potential role for Virtual Reality (VR) software use (Penn, Rose & Johnson, 2008). VR offers a way of creating more realistic ‘real world’ activities within the clinic or laboratory in which task demands can be made replicable and performance can be automatically recorded (Zhang et al., 2003; Parsons, 2015). The potential use within neuropsychological assessment and rehabilitation has been recognised (Schultheis & Rizzo, 2001; Rizzo et al., 2004a), including simulating situations and tasks that people experience in their daily lives, such as shopping (Lo Priore et al., 2003) and driving (Liu et al., 1999), within safe, controlled and standardised formats (Morris, 2005).

Nevertheless, there have been few examples of VR procedures developed to test EF. An early example is the VR ‘Bungalow Task’ (Morris, Kotsitsa, Bramham, Brooks & Rose, 2002) which has been shown to be sensitive to planning impairments in individuals with damage to PFC (see also Sweeney, Kersel, Morris, Manly & Evans, 2010). Participants are required to take on the role of a ‘removal person,’ moving around the rooms of a building to find specified furniture to be removed. Furniture had to be chosen appropriately for the rooms of the house and collected in a particular order, according to its category. Time-based and event-based tests of PM were embedded in the task. A frontal lobe lesion (FLL) group visited fewer rooms and showed less efficient strategies, increased rule breaks and impairments in PM compared to controls. There is also promising evidence that VR assessments can accurately identify EF impairments in individuals with acquired brain injury (ABI), rather than FLL specifically (Sweeney et al., 2010).

Another VR task for measuring EF is the Jansari assessment of Executive Functions (JEF[®]). In this task, participants take on the role of an office worker whose

primary objective is to organise and prepare for a meeting and the various subtasks successfully mimic everyday multi-tasking requirements. This procedure has the advantage that it has been validated with different populations and it appears to be sensitive at detecting the impact of chemicals on EF (Montgomery, Hatton, Fisk, Ogden & Jansari, 2010; Montgomery, Ashmore & Jansari, 2011; Montgomery, Seddon, Fisk, Murphy & Jansari, 2012; Jansari et al., 2013; Soar, Chapman, Lavan, Jansari & Turner, 2016). In terms of concurrent validation, Renison, Ponsford, Testa and Jansari (2008) compared individuals with ABI and control participants on their performance on the task with other measures of EF, including the Modified Six Elements Test and the Zoo Map Test from the BADS, finding comparable sensitivity. Jansari et al., (2014) also compared the performance of 17 individuals with acquired brain injury (ABI) with that of 30 healthy controls across eight JEF[®] EF constructs, namely: planning, prioritisation, selection, creative thinking, adaptive thinking, action-based PM, event-based PM, and time-based PM. The task differentiated between individuals with ABI and controls on each construct as well as on overall performance. In this study, JEF[®] was better able to detect more complex aspects of executive dysfunction than the other EF measures used (Jansari et al., 2014). The task may further have merit in being used to test rehabilitation strategies or pharmacological interventions that are used with individuals with ABI (Yesavage et al., 2007; Hosenbocus & Chahal., 2013).

In the Jansari et al., (2014) study, the ABI participants had widespread and heterogeneous lesions, including brain damage ranging from right fronto-parietal to frontal, temporal, anterior, and occipital areas, also consisting of a range of aetiologies including head injuries. Whilst such participants reflect the range of patients likely to be encountered in a neurorehabilitation setting, there are advantages in validating a task in groups of individuals who have more circumscribed brain lesions likely to affect EF. Studying the effects of focal brain lesions is a way of testing ‘proof of principle’ relating to specific tasks when considering the anatomical and functional relationships of particular brain areas. Additionally measured deficits can be shown to be more specific to the intended function, rather than a consequence of general under-function. Additionally, neurosurgical mapping techniques with focal lesion patients can demonstrate which neurocognitive systems are involved in task performance (e.g. Manes et al., 2002; Hornak et al., 2004; Pullen, Morris, Kerr, Bullock & Selway, 2006; Bramham et al., 2009; Lovstad et al., 2012).

In the present study, individuals with specific unilateral and bilateral surgical excisions for tumours in the frontal lobes were tested on JEF[®], and their performance was compared with that of healthy controls. The primary objective of the current study was to determine whether a VR test of multitasking would detect the difficulties in EF that are frequently reported by and/or observed in individuals with circumscribed FLL in everyday life. Comparisons were made with non-VR EF measures and questionnaires that focused on real-life EF dysfunction.

METHODS

Participants

Nineteen individuals with focal frontal lobe (FLL) lesions were recruited from the joint neuro-oncology clinic at King's College Hospital, London. Only individuals with lesions exclusive to the PFC were selected. The exclusion criteria included the following: the presence of additional neurological conditions, autism spectrum disorder or attention deficit hyperactivity disorder, psychiatric conditions, a history of dependency on drugs or alcohol, language impairment, hearing or visual difficulties. The test procedures all involved verbal instructions in English, and as a consequence, potential participants who were not fluent in English were also excluded. During the first testing session, participants were screened on measures of current intellectual functioning and only those who had had IQ scores >70 were included. They were tested at least six months post-surgery (M: 38.52, SD: 36.09, range: 6-106) to reduce acute post-operative effects on cognitive functioning. All lived independently in the community.

Nineteen healthy controls were recruited, group matched with the FLL group for age, years of education, estimated pre-morbid IQ and gender (FLL: 10F, 9M, controls: 10F, 9M, see Table 1). There was a statistically significant difference between groups on Full-Scale IQ measured using the abbreviated two-subtest version of the Wechsler Abbreviated Scale of Intelligence (WASI-II; Wechsler, 2011).

Participants gave written informed consent and the study was approved by a local research governance committee and the London Bridge National Research Ethics Service Committee.

(Table 1 about here)

The method used by Rowe, Bullock, Polkey & Morris (2001) was adopted to classify lesion areas (see Table 2). These were verified by the neurosurgeon by inspection of Magnetic Resonance Imaging (MRI) or Computerised Tomography (CT) scans and neuroradiological reports defining brain involvement in terms of Brodmann areas (Brodmann, 1909). Seven individuals had right frontal lobe lesions, nine had left frontal lobe lesions and three had bilateral lesions. Brodmann encroachment was amalgamated into three main PFC regions, (see Table 2), defined anatomically as dorsolateral (Brodmann areas 44, 45 and 46), medial (Brodmann areas 8, 9, 24, 25 and 32) and orbitofrontal regions (Brodmann areas 10, 11, 12 and 47).

(Table 2 about here)

Measures

A battery of standardised tests was administered to all participants. These were chosen to measure intellectual function, memory and EF. The Logical Memory and Visual Reproduction subtests of the Wechsler Memory Scale- Fourth UK Edition (WMS-IV; Wechsler, 2009) were given as measures of auditory memory and visual memory respectively, with immediate and delayed memory tested. Measures of working memory consisted of the Digit Span subtest of the Wechsler Memory Scale-Third UK Edition (WMS-III; Wechsler, 1997) and the Spatial Span subtest of the Wechsler Memory Scale-Third UK Edition (WMS-III; Wechsler, 1997). The Sustained Attention to Response Task (SART; Robertson, Manly, Andrade, Baddeley, and Yiend, 1997) measured attention, administered using a laptop computer.

In addition, both groups were tested on a battery of frequently used clinical tests of EF tests, namely the Trail Making Test Part A and Part B (TMT; Army Individual Test Battery 1944; Reitan, 1992), the Hayling Sentence Completion Test and the Brixton Spatial Anticipation Test (Burgess and Shallice, 1997) and verbal fluency FAS measures from the Delis-Kaplan Executive Function System (DKEFS; Delis, Kaplan & Kramer, 2001).

Questionnaires

Two questionnaires that measure EF and are used widely in brain injury populations were administered to all participants. This includes the Frontal Systems Behaviour

Scale (FrSBe, Grace & Malloy, 2001), a 46-item rating scale that provides a brief, reliable, and valid measure of three frontal systems behavioural syndromes: apathy, disinhibition, and executive dysfunction. The FrSBe quantifies behavioural changes over time by including both baseline (retrospective) and current assessments of behaviour, including apathy, disinhibition and executive function. Healthy controls were asked to only complete current ratings. In addition, the study used a revised and extended version of the Dysexecutive Questionnaire (DEX; Wilson, Alderman, Burgess, Emslie & Evans, 1997) developed by Simblett, Ring and Bateman (2016). Total scores were calculated for each of the four domains: Emotional-Behavioural Self-regulation (maximum score /36), Activation (maximum score /32), Metacognition (maximum score /32) and Executive Cognition (maximum score/ 40). Higher scores indicated greater difficulties.

Measures of apathy, anxiety and depression were also used, since such difficulties are common in people with tumours involving the frontal lobe. For apathy, the Apathy Evaluation Scale (AES) was used, an 18-item scale developed by Marin (1991) specifically for use in populations with brain-related pathology. The AES evaluates the overt behavioural, cognitive, and emotional aspects of goal-directed behaviour (Marin, 1991). Each AES form yielded a total score, with higher scores indicating the presence of a greater degree of apathy. Cut-off scores of 41 were used as stated in the AES guidelines. The Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith, 1983) was used as a screening measure of anxiety and depression, with the two subscales each scored in the ranges of 0-21: Scores of 0-7 are considered normal, 8-10 borderline, and above 11 clinically significant.

The Jansari assessment of Executive Functions (JEF©)

This task was presented in a desktop VR environment, on a laptop, with the systems unit using Microsoft Visual Basic and the 3D add-on software 3d State (<http://www.3dstate.co.uk/wordpress/>) as a platform for the specific software (see Figures 1-3 for visual representations). It was administered following the standard procedure outlined in the manual (Jansari, unpublished).

JEF© is set in an office environment and the participant is asked to imagine that they are starting their first day as an office worker. A scenario is presented whereby their manager has been called away so will not be able to oversee their work, but has left the participant a list of jobs that they need to do to prepare for a meeting. There are

two rooms in the environment, an office and a meeting room. A corridor links these rooms and the participant can move freely between them. Realistic tasks that can be found in an average office environment are chosen for eight different cognitive constructs: planning, prioritisation, selection, creative thinking, adaptive thinking, action-based PM, event-based PM, and time-based PM. Tasks were designed to be ambiguous and have multiple solutions, to mimic real-life situations. The three main task categories related to a 'meeting', doing 'the post', and additional time-based tasks. A printed scenario sheet, the Manager's Tasks for Completion, and all relevant documents (post diary, list of the post to be sent, agenda topics, My Notes For Manager and plan of action) were provided to the participant, outside the virtual environment. They remained next to the computer throughout the assessment for participants. Participants were allowed to write on the material; for example, they could add to the notes for the manager or tick off the tasks on their plan of action, and use this as an aid to reduce the likelihood of errors being made due to failures of retrospective memory.

Before starting the task, the participant practised manoeuvring within the virtual environment using the arrow keys on a standard computer keypad. Objects were picked up by clicking the computer mouse. At the beginning, the task scenario was read to the participant from a script. After reading the Manager's Tasks for Completion, participants were required to construct a plan of action in their own time, before the VR component of the assessment formally commenced. The experimenter directed participants to the printed materials if they had task-specific questions. In addition, various PM tasks were built into the procedure. Specifically, individuals were handed a number of memoranda throughout the assessment, which required them to complete additional tasks at set points later in time. The responsibility for planning the overall task was given to participants with no clues as to solutions or courses of action. They were given 40 minutes to complete the list of tasks in time for the beginning of the meeting. If they exceeded this, they were allowed to continue and their total time taken was recorded, but not included in the overall score. The start time and the meeting time were both written down and participants had a digital clock in front of them so that they could monitor the time. The experimenter observed the assessment and filled out the score-sheet while participants were completing the task.

(Figures 1-3 about here)

RESULTS

Background neuropsychological measures

The FLL group had significantly worse immediate ($t(36) = 2.7, p < .01, d = -.87$) and delayed ($t(36) = 2.6, p < .02, d = -0.84$) visual recall compared to controls. The groups had comparable performance for visual recognition memory $t(36) = 1.3, p = 0.18$, digit span $t(36) = 0.87, p = 0.38$, spatial span $t(36) = 0.87, p = 0.38$ and on the SART (errors of commission $t(36) = 0.95, p = 0.34$, omission $t(36) = 1.3, p = 0.19$ and mean reaction time $t(36) = 0.45, p = 0.65$).

Non-VR EF measures

The non-VR EF measure results are shown in Table 3. The FLL group were significantly slower than the controls on the comparison Trail Making A test, but not on the Trail Making B, which measures mental flexibility. There were also no significant differences between groups on the Hayling and the Brixton. There was a marginally significant difference between groups in the total number of items generated on verbal fluency. Analyses were also conducted using an ANCOVA to covary for the significant difference in IQ between groups; there were no significant differences across any of the EF measures when the effect of FSIQ was covaried.

(Table 3 about here)

These findings suggest that with the exception of Trail Making Test A, the standard measures of EF were unable to distinguish between the FLL and control groups.

Questionnaires

Questionnaires were completed by 16 individuals in the FLL group and 19 individuals in the control group through self-report. Individuals with FLL reported significantly higher symptoms on the FrSBe as rated currently, compared to before their surgery $t(13) = 2.28, p < .041, d = -0.47$ (after: M: 56.23, SD: 16.94, before: M: 48.7, SD: 11.17).

A comparison of the FLL and control groups revealed no significant differences between groups on the FrSBE $t(27)=1.20$, $p=.24$ and the four scales of the DEX: emotional behavioural self-regulation scale $t(33)=1.48$, $p=0.14$, activation scale $t(33)=1.16$, $p=0.25$, metacognition scale $t(33)=1.72$, $p=0.95$ and executive cognition scale $t(33)=1.78$, $p=0.83$. There were no between group differences on the AES $t(32)=.44$, $p=0.66$ and on HADS anxiety $t(33)=1.68$, $p=0.10$ and HADS depression scales $t(33)=1.68$, $p=0.10$.

The Jansari assessment of Executive Functions

All tasks were scored on a three-point scale: 0 for failure, 1 for a partial or non-optimal completion and 2 for satisfactory completion. Construct scores were created by amalgamation of tasks scores with some constructs involving only one task and others including two; to allow comparisons, a percentage score was calculated for each construct. An overall percentage score was obtained by averaging the individual construct scores. In all, nine scores were derived for each participant, eight for the individual constructs and one for average performance. A between subjects ANOVA demonstrated that the overall score of the FLL group was significantly lower than that of the control group, with the effect size of this difference being considered large according to Cohen's (1992) guidelines, $F(2, 37)=17.21$, $p<.001$, $\eta_p^2=3.2$ (see Figure 4). Given the significant difference in FSIQ between groups, an ANCOVA was conducted to covary for the effect of FSIQ between groups. However, the difference remained significant $F(2, 37)=9.89$, $p<.003$, $\eta_p^2=.22$ (group), $F(2, 37)=13.17$, $p<.001$, $\eta_p^2=.27$ (FSIQ).

(Figure 4 about here)

Comparisons of the eight individual constructs were conducted using non-parametric analyses, as the Shapiro-Wilk test demonstrated that the data were not normally distributed. There was a significant difference between groups on planning: $U(38)=254$, $p<.03$, creative thinking: $U(38)=252$, $p<.03$, adaptive thinking: $U(38)=266.5$, $p<.01$, event-based PM: $U(38)=272.5$, $p<.006$, and time-based PM: $U(38)=276.5$, $p<.004$ (see Table 4 for effect sizes). There were no significant differences between groups for prioritisation, selection, and action-based PM.

(Table 4 about here)

Analysis of individual performance

To assess individual performance within the FLL group relative to the control group, percentiles were created for each construct using the control group data (see Table 5).

(Table 5 about here)

Individuals in the FLL group with scores below the 5th or between the 6th and 10th percentile were then identified for each construct (see Table 6), and their frequencies examined. For the 5th percentile cut-off, the constructs upon which the greatest number of individuals within the FLL group showed impairment were adaptive thinking (n=6), followed by creative thinking (n=5), action-based PM (n=5), time-based PM (n=4) and prioritisation (n=4). It should be noted that some individuals in the control group also had impaired scores for two constructs: creative thinking (n=3) and action-based PM (n=5). Performance across the constructs was variable. None of the FLL individuals were impaired in all domains. Three out of nineteen individuals had impaired average scores. Five individuals each had impaired performance on none, one, and two constructs. This was followed by three constructs (n=1), or four constructs (n=3). When looking at the frequencies of FLL individuals with scores in the 6-10th percentile range, the average score had the greatest number (n=12), followed by adaptive thinking (n=6), prioritisation (n=6), creative thinking (n=5) and action-based PM (n=5). Six individuals in the FLL group had scores in this range on three constructs, this was followed by two constructs (n=3), five constructs (n=3), four constructs (n=1) and one construct (n=1).

(Table 6 about here)

Executive Function composite

The overall task score may be better able to identify group differences because it acts as a composite for many different individual task constructs including, for example, planning, prioritisation and prospective memory. The EF tasks used in this study measure fewer constructs than the JEF[®], for example, the Hayling measures inhibition and response initiation, so the tasks may not be directly comparable to the overall JEF[®] score. In order to address this difference in measurement, an EF composite measure was created from the individual EF measures (Trails A percentile, Trails B percentile, Brixton scaled, Hayling scaled and FAS percentile) and this EF composite was

compared with the overall score. To calculate the composite score, each individual EF measure was converted into a z-score using the mean and standard deviation of the healthy control group to ensure that all measures were on the same scale. An inter-item total correlation was carried out to ensure each z-score converted EF measure was a suitable variable to be included in the composite measure. An inter-item correlation cut-off of .03 was used to justify the inclusion of each measure and each item was above .05 (Streiner & Norman, 2008). Cronbach's alpha was .66 and this value did not change considerably when each measure was removed. Therefore, all five measures were included in the composite.

Independent t-tests demonstrated a significant difference between groups on the composite non-VR EF z-score measure $t(35)=2.05$, $p<.04$, $d=-0.66$ (FLL: M: $-.46$, SD: 1.30 , control: M: $.00$, SD: 1.00) and a significant difference on the average JEF[®] z-score $t(36)=4.14$, $p<.001$, $d=-1.34$ (FLL: M: -1.56 , SD: 1.30 , control: M: $.00$, SD: 1.0). For the FLL group, a paired t-test showed that the overall JEF[®] z-score was significantly lower than the EF composite z-score $t(18)=3.48$, $p<.003$, $d=-0.92$ (FLL composite: M: $-.46$, SD: 1.30 ; FLL JEF[®]: M: -1.56 , SD: 1.30) indicating that the JEF[®] is better at differentiating between groups compared to the EF composite.

Sensitivity and specificity analysis

The ROC curve graphically displays the trade-off between sensitivity and specificity and is useful in assigning the best cut-offs for clinical use (Florkowski, 2008). The area under the curve (AUC) determines the inherent ability of a test to discriminate between "healthy and diseased populations" (Hajian-Tilaki, 2013). In a Receiver Operating Characteristic (ROC) curve analysis applied to the overall JEF[®] score, the AUC was .83 and a cut-off value of 66.15 was determined. This resulted in 73.7% sensitivity and 89.5% specificity for the average score. This indicated that 73.7% of FLL individuals were correctly classified and 10.5% controls were incorrectly classified, which suggests good sensitivity and specificity (Harris & Taylor, 2014).

(Figure 5 about here)

Lesion analyses

Supplementary analyses were conducted to investigate the effects of laterality and location of lesions within the frontal lobe group in terms of JEF[®] performance and the non-VR EF measures. The method used by Rowe et al., (2001) was adopted, where individuals who had an operation in a specific location were compared to the rest of the sample who did not have an operation in this region. For laterality analyses, unilateral left (n = 9) were compared with unilateral right hemisphere lesions (n = 7) (this excluded the three bilateral lesion individuals); for lesion location analyses dorsolateral, non-medial lesions (n = 4) were compared with non-dorsolateral, medial lesions (n = 15) and finally, orbitofrontal lesions (n=6) were compared with non-orbitofrontal lesions (n=13). No significant effects of laterality or lesion location were found on JEF[®] or non-VR EF measures.

DISCUSSION

A comparison between individuals with FLL and matched controls on an ecologically valid VR measure of EF, namely JEF[®], demonstrated an overall group difference. The FLL group were impaired on five out of eight possible task constructs: planning, creative thinking, adaptive thinking, event-based PM and time-based PM, with no significant difference on prioritisation, selection, and action-based PM. In this group of people with circumscribed FLL lesions, the VR measure was shown to be sensitive to EF deficits whilst frequently used clinical tests of EF were not. In the study by Jansari (2014), the deficits were found in more constructs, which may reflect the more specific lesions and less generalised effect in our study. In the current study, the groups were matched on age, years of education, and premorbid IQ, whereas in the previous study, the groups were only matched on age and premorbid IQ. The ABI group tested by Jansari et al., (2014) used a mixed clinical sample, including participants with injuries of various aetiologies including stroke and traumatic brain injury, which are associated with larger lesions with more diffuse damage. They were thus more likely to have additional cognitive difficulties, which would exacerbate group differences in JEF[®] performance.

An analysis of individual performance in the FLL group using control group percentiles demonstrated that not all individuals were impaired on the same constructs. This finding of heterogeneity of performance was also found in Jansari et al.,'s (2014) study and reflects the fact that individual EF tasks in general tend to have low

correlations with one another, including when measured using ecologically valid tasks (Burgess, Simons, Coates & Shannon, 2005).

There were no group differences on the questionnaires and no discrepancies between the FLL self and other report measures. This finding is consistent with other research. Gregg et al., (2014) compared frontal and non-frontal tumour groups on the FrSBe and found no differences between self and informant reports within their frontal group. In addition, Lengenfelder et al., (2015) found no significant differences between individuals with Traumatic Brain Injury (TBI) and family members' reports for any of the FrSBe subscales. The FLL group reported significantly higher post-injury difficulties as reflected in the overall scores of the FrSBe relative to pre-injury scores. This finding also replicates other research studies with similar populations (Gregg et al., 2014; Lengenfelder et al., 2015). The lack of significant difference between FLL and control groups on any of the questionnaire measures is notable, with little research directly comparing questionnaire responses from individuals with frontal lobe lesions and healthy controls. Grace, Stout and Malloy (1999) found significantly more 'frontal behaviour' in frontal lesion groups than controls. The lack of sensitivity in the current study might reflect the fact that we recruited subjects from an outpatient neuro-oncology department where patients attended for routine oncological follow up, rather than because they had cognitive or behavioural difficulties following their surgery. In other studies, individuals with FLL may be recruited from inpatient and rehabilitation settings where these difficulties may be more prominent. Our findings may therefore indicate that the more subtle behaviour changes are not picked up in such patients by questioning but can be measured using VR ecological valid procedures.

The FLL and controls are distinguished on JEF[®] average performance and across five individual constructs. In contrast, the majority of EF measures did not distinguish between groups. These findings are congruent with a number of other studies in the field demonstrating a group difference on ecologically valid measures and comparable performance on non-VR well-used EF measures (Eslinger & Damasio, 1985; Shallice & Burgess, 1991; Burgess et al., 1998; 2006).

As there is a composite JEF[®] score sampling various executive domains, a composite measure was created for the individual non-VR EF tasks in order to provide a direct comparison with the VR measure. There was a significant difference in composite EF scores between FLL and control groups. A within-group analysis

demonstrated the FLL group had poorer overall JEF[®] z-scores than EF composite z-scores. However, just as for previously used EF measures, whilst a group finding supports use of a composite score, heterogeneity between individuals on what particular measures show deficits suggest consideration of individual scores.

The action-based PM was the most difficult task for those in the FLL group, and the second most-difficult task for those in the control group, with both groups achieving scores of 30-40%. There is little research on action-based PM. It is considered easier than time and event-based PM because it does not require the interruption of ongoing activity (Kvavilashvili & Ellis, 1996). Shum, Valentine and Cutmore (1999) showed that individuals with TBI and controls had better performance on action-based than time and event-based PM tasks. However, Brewer et al., (2011) found that action-based performance was more impaired than comparable event-based conditions in healthy volunteers. One potential contributor to the relatively weak performance on the JEF[®] action-based PM tasks is that this construct differs from the others, as it requires two steps. The participant has to carry out an action and then write down that it had been completed rather than just reorganise the post. The result on action-based PM was not the focus of the current study, yet it raises interesting questions for further research.

Our results indicate JEF[®] is suitable for use with individuals with FLL, with all participants able to follow the basic procedures and navigate around the office scenario. The PFC group was challenged by the VR procedure and this may account for the task sensitivity. Marcotte and colleagues (2010) noted the difficulty in developing measures reflective of daily functioning in a manner that is “sufficiently challenging to provide a distribution of functioning across ‘normal’ individuals” (p24) such that ceiling and floor effects are avoided. JEF[®] was found to be appropriate for the range of control participants and patients used in the study and was not subject to such effects.

In the current study, supplementary analyses within the frontal lobe group indicated that there were no laterality and lesion location effects. The sample size and range of lesions mean it was not possible to make any firm conclusions on these matters. The majority of individuals recruited in the FLL group had parafalcine tumours, which resulted in medial lesions. Further exploration with a bigger and more varied sample of individuals with FLL needs to be conducted. Additionally, studies with larger sample sizes of individuals with FLL would also answer questions

regarding how performance on the JEF[®] fits with theoretical accounts regarding fractionation of the EF system (Stuss and Alexander, 2007).

Conclusions and implications

The study demonstrated that individuals with FLL did not differ significantly from matched controls in their self-reported difficulties with executive functioning, or on performance on non-VR EF measures. However, the FLL group were impaired relative to controls on their JEF[®] performance. The present study expands on previous research, providing support for the use of VR ecologically-valid measures that discriminate between individuals with FLL and controls. The findings suggest the task measures EF dysfunction more specifically related to frontal function. The task highlights specific cognitive constructs that individuals have difficulty with, for example, prospective memory, which can be directly targeted in interventions. An important implication is that one should not presume that VR and non-VR measures of EF capture the same level of underlying process or neural substrate. Both measures are useful and valuable and in combination they provide a more complete picture during clinical assessment.

ACKNOWLEDGEMENTS

The authors wish to thank the neuro-oncology team at Kings College Hospital for their support in identifying patients, Dr Sara Simblett and Dr Andrew Bateman for allowing us to use the DEX-R questionnaire, Dr Daniel Stahl for his statistical advice and the participants who generously gave their time and enthusiasm to take part in this research. The author(s) reported no potential conflict of interest. This work was supported by the Doctorate in Clinical Psychology program at the Institute of Psychology, Psychiatry and Neuroscience, King's College London.

REFERENCES

- Anderson, C.V., Bigler, E.D., & Blatter, D.D. (1995). Frontal lobe lesions, diffuse damage, and neuropsychological functioning in traumatic brain injured patients. *Journal of Clinical and Experimental Neuropsychology*, 17, 900-908.
- Aron, A.R., Monsell, S., Sahakian, B.J., & Robbins, T.W. (2004). A component analysis of task-switching deficits associated with lesions of left and right frontal cortex. *Brain*, 127, 1561-73.
- Army, U. (1944). *Army Individual Test Battery. Manual of Directions and Scoring*.
- Baddeley, A., (1986). *Working Memory*. Oxford University Press, Oxford.
- Bechara, A., Damasio, H., & Damasio, A. (2000). Emotion, Decision Making and the Orbitofrontal cortex. *Cerebral Cortex*, 10 (3), 295-307.
- Bisiacchi, P. S., Sgaramella, T. M., & Farinello, C. (1998). Planning strategies and control mechanisms: evidence from closed head injury and aging. *Brain and Cognition*, 37, 113–116.
- Bonelli, R.M., & Cummings, J.L. (2007). Frontal-subcortical circuitry and behavior. *Dialogues in Clinical Neuroscience*, 9, 141-151.
- Bramham, J., Morris., R.G., Hornak, J., Bullock, P., & Polkey, C.E. (2009). Social and emotional functioning following bilateral and unilateral neurosurgical prefrontal cortex lesions. *Journal of Neuropsychology*, 3, 125-143.
- Brewer, G.A., Marsh, R.I., Clark-Foos, A., Thadeus Meeks, J., Cook, G.I., & Hicks, J.L. (2011). A comparison of activity-based to event-based prospective memory. *Applied Cognitive Psychology*, 25, 632-640.
- Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt auf Grund des Zellenbaues* (Barth, Leipzig); English translation available in Garey, L. J. Brodmann's Localization in the Cerebral Cortex (Smith Gordon, London, 1994).
- Burgess, P. W., & Shallice, T. (1997). *The Hayling and Brixton tests*. Thurston. Suffolk: Thames Valley Test Company.
- Burgess, P., Alderman, N., Evans, J., Emslie, H., & Wilson, B. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*, 4, 547-558.
- Burgess, P.W., Veitch, E., Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, 38, 848-863.

Burgess, P.W., Simons, J.S., Coates, L.M.A., & Channon, S. (2005). The search for specific planning processes. In R. Morris & G. Ward (Eds.), *The Cognitive Psychology of Planning* (pp. 199-227). Hove, U.K.: Psychology Press.

Burgess, P.W., Simons, J.S., Dumontheil I., & Gilbert, S.J. (2005b). The gateway hypothesis of rostral Pre Frontal Cortex (area 10) function. In J.Duncan, P.McLeod & L.Phillips (Eds.), *Measuring the Mind: Speed, Control and Age* (pp. 215-246). Oxford University Press.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Coates, L. M. A., Dawson, D. R., Anderson, N. D., Gilnert, S. J., Dumontheil, I., & Channon, S. (2006). The case for the development and use of “ecologically valid” measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, 12, 194-209.

Burgess, P.W., Alderman, N., Volle, E., Benoit, R.G., & Gilbert, S.J. (2009). Mesulam's frontal lobe mystery re-examined. *Restorative Neurology and Neuroscience*, 27, 493-506.

Crépeau, F., Belleville, S., & Duchesne, G. (1996). Disorganisation of Behavior in a Multiple Subgoals Scheduling Task Following Traumatic Brain Injury. *Brain and Cognition*, 32, 266-268.

Dawkins, N., Cloherty, M.E., Gracey, F., & Evans, J.J. (2006). The factor structure of the Hospital Anxiety and Depression Scale in acquired brain injury. *Brain Injury*, 20 (12), 1235-9.

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001a). *Delis-Kaplan Executive Function System (D-KEFS)*. San Antonio, TX: The Psychological Corporation.

Dias, R., Robbins, T.W., & Roberts, A.C. (1996). Dissociation in prefrontal cortex of affective and attentional shifts. *Nature*, 7, 69-72.

Elliott, R. (2003). Executive functions and their disorders. Imaging in clinical neuroscience. *British Medical Bulletin*. 65 (1), 49-59.

Ellis, J. (1996). Prospective memory or the realization of delayed intentions: A conceptual framework for research. In M. Brandimonte, G. O. Einstein, & M. A. McDaniel (Eds.), *Prospective memory: Theory and applications* (pp 1-22). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Ellis, J., & Freeman, J. E. (2008). Ten years on: Realizing delayed intentions. In M. Kliegel, M. A., McDaniel, & G. O. Einstein (Eds.), *Prospective memory: Cognitive, Neuroscience, Developmental, and Applied Perspectives* (pp. 1-27). New York: Lawrence Erlbaum Associates.

Eslinger, P.J., & Damasio, A.R. (1985). Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology*, 35, 1731-1741.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39 (2), 175-191.

Florkowski, C.M. (2008). Sensitivity, Specificity, Receiver-Operating Characteristic (ROC) Curves and Likelihood Ratios: Communicating the performance of Diagnostic Tests. *The Clinical Biochemist Reviews*, 29, S83-S87.

Funahashi, S. (2001). Neuronal mechanisms of executive control by the prefrontal cortex. *Neuroscience Research*, 39, 147-165.

Gioia, G.A., Isquith, P.K., Guy, S.C., & Kenworthy, L. (2000). *Behavior Rating Inventory of Executive Function (BRIEF): professional manual*. Lutz, FL: Psychological Assessment Resources.

Goel, V., & Dolan, R.J. (2003). Reciprocal neural response within lateral and ventral medial prefrontal cortex during hot and cold reasoning. *Neuroimage*, 20, 2314-21.

Goldman-Rakic, P.S. (1996). The prefrontal landscape: implications of functional architecture for understanding human mentation and the central executive. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences*, 29, 1445-53.

Goldstein, L.H., Bernard, S., Fenwick, P.B., Burgess, P.W., & McNeill, J. (1993). Unilateral frontal lobectomy can produce strategy application disorder. *Journal of Neurology, Neurosurgery and Psychiatry*, 56 (3), 274-276.

Grace, J., Stout, J. C., & Malloy, P. F. (1999). Assessing frontal lobe behavioral syndromes with the frontal lobe personality scale. *Assessment*, 6, 269–284.

Grace, J., & Malloy, P.F. (2001). *FrSBe. Frontal systems Behavior Scale: Professional manual*. Lutz: FL: Psychological Assessment Resources, Inc.

Gregg, N., Arber, A., Ashkan, K., Brazil, L., Bhangoo, R., Beaney, R., Gullan, R., Hurwitz, V., Costello, A., & Yaguez, L. (2014). Neurobehavioural changes in patients following brain tumour: patients and relatives perspective. *Support Care Cancer*, 22, 2965-72.

Hajian-Tilaki, K. (2013). Receiver Operating Characteristic (ROC) Curve Analysis for Medical Diagnostic Test Evaluation. *Caspian Journal of Internal Medicine*, 4 (2), 627-635.

Harris, M., & Taylor, M. (2014). *Medical statistics made easy*. 3rd edition. Scion Publishing Ltd. Ian Potter Library.

Heaton, R.K. (1981). *The Wisconsin Card Sorting Test Manual*. Odessa, FL: Psychological Assessment Resources.

- Hosenbocus, S., & Chalhal, R. (2012). A review of executive function deficits and pharmacological management in children and adolescents. *Journal of the Canadian Academy of Child and Adolescent Psychiatry*, 21 (3), 223-229.
- Hornak, J., O'Doherty, J., Bramham, J., Rolls, E.T., Morris, R.G., Bullock, P.R., & Polkey, C.E. (2004). Reward related reversal learning after surgical excisions in orbito-frontal or dorsolateral prefrontal cortex in humans. *Journal of Cognitive Neuroscience*, 16, 3, 463-478.
- Hsu, W., Zanto, T.P., Anguera, J.A., Lin, Y., & Gazzaley, A. (2015). Delayed enhancement of multitasking performance: Effects of anodal transcranial direct current stimulation on the prefrontal cortex. *Cortex*, 69, 175-185.
- Jansari, A.S., Froggatt, D., Edginton, T., & Dawkins, L. (2013). Investigating the impact of nicotine on executive functions using a novel virtual reality assessment. *Addiction*, 108 (5), 977-84.
- Jansari, A., Devlin, A., Agnew, R., Akesson, K., Murphy, L., & Leadbetter, T. (2014). Ecological Assessment of Executive Functions: A New Virtual Reality Paradigm. *Brain Impairment*, 15 (2), 71-87.
- Jansari, A. (2014). JEF©-The Jansari assessment of Executive Functions manual. Unpublished.
- Khalighinejad, N., Di Costa, S., & Haggard, S. (2016). Endogenous Action Selection Processes in Dorsolateral Prefrontal Cortex Contribute to Sense of Agency: A Meta-Analysis of tDCS Studies of 'Intentional Binding'. *Brain Stimulation*, 9 (3), 372-379.
- Kvavilashvili, L., & Ellis, J. (1996). Varieties of intention: Some distinctions and classifications. In M. Brandimonte, G.O. Einstein & M.A. McDaniel (Eds.), *Prospective Memory: Theory and Applications* (pp. 23-52). Mahwah, NJ: Lawrence Erlbaum.
- Lengenfelder, J., Arjunman, A., Chiaravalloti, N., Smith, A., & Deluca, J. (2015). Assessing frontal behavioral syndromes and cognitive functions in traumatic brain injury. *Applied Neuropsychology*, 22, 7-15.
- Levy, R., & Dubois, B. (2006). Apathy and the functional anatomy of the prefrontal cortex-basal ganglia circuits. *Cerebral Cortex*, 16 (7), 916-28.
- Lezak M., Howieson D., Bigler E., & Tranel D. (2012). *Neuropsychological Assessment*. New York, NY: Oxford University Press.
- Liu, L., Miyazaki, M., & Watson, B. (1999). Norms and validity of the DriVR: A virtual reality driving assessment for persons with head injuries. *Cyberpsychology and Behavior*, 2, 53-67.
- Lo Priore, C., Castelnovo, G., Liccione, D., & Liccione, D. (2003). Experience with V-STORE: considerations on presence in virtual environments for effective

neuropsychological rehabilitation of executive functions. *Cyberpsychology and Behavior*, 6, 281-7.

Lovstad, M., Funderud, I., Endestad, T., Due-Tønnessen, P., Meling, T.R., Lindgren, M., Knight, R.T., & Solbakk, A.K. (2012). Executive functions after orbital or lateral prefrontal lesions: Neuropsychological profiles and self-reported executive functions in everyday living. *Brain Injury*, 26, 1586-1598.

Manes, F., Sahakian, B., Clark, L., Rogers, R., Antoun, N., Aitken, M., & Robbins, T. (2002). Decision-making processes following damage to the prefrontal cortex. *Brain*, 125 (3) 624-39.

Manly, T., Hawkins, K., Evans, J., Woldt, R., & Robertson, I.H. (2002). Rehabilitation of executive function: facilitation of effective goal management on complex tasks using periodic auditory alerts. *Neuropsychologia*, 40, 271-281.

Marin, R.S., Biedrzycki, R.C., & Firinciogullari, S. (1991). Reliability and validity of the Apathy Evaluation Scale. *Psychiatry Research*, 38, 143-162.

Miotto, E. C., & Morris, R. G. (1998). Virtual planning in patients with frontal lobe lesions. *Cortex*, 34 (5), 639-657.

Miyake A., Friedman N. P., Emerson M. J., Witzki A. H., Howerter A., & Wager T. D. (2000). The unity and diversity of executive functions and their contributions to complex “frontal lobe” tasks: a latent variable analysis. *Cognitive Psychology*, 41, 49–100.

Monchi, O., Petrides, M., Petre, V., Worsley, K., & Dagher, A. (2001). Wisconsin Card Sorting revisited: distinct neural circuits participating in different stages of the task identified by event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, 21, 7733–7741.

Montgomery, C., Hatton, N.P., Fisk, J.E., Ogden, R.S., & Jansari, A.S. (2010). Assessing the functional significance of ecstasy-related memory deficits using a virtual paradigm. *Human Psychopharmacology: Clinical and Experimental*, 25(4), 318–325.

Montgomery, C., Ashmore, K., & Jansari, A. (2011). The effects of a modest dose of alcohol on executive functioning and prospective memory. *Human Psychopharmacology: Clinical and Experimental*, 26, 208–215

Montgomery, C., Seddon, A.L., Fisk, J.E., Murphy, P.N., & Jansari, A.S. (2012). Cannabis-related deficits in real-world memory. *Human Psychopharmacology*, 27 (2), 217–225.

Morris, R.G., Kotitsa, M., Bramham, J., Brooks, B.M., & Rose, F.D. (2002). Virtual reality investigation of strategy formation, rule breaking and prospective memory in patients with focal prefrontal neurosurgical lesions. *ICDVRAT*, 4, 101-108.

Morris RG (2005). Virtual reality in brain damage rehabilitation: Commentary. *Cyber-Psychology and Behaviour*, 8 (3), 267-268.

Murray, E. A., Bussey, T. J., & Wise, S. P. (2000). Role of prefrontal cortex in a network for arbitrary visuomotor mapping. *Experimental Brain Research*, 133 (1), 114-129.

Neulinger, K., Oram, J., Tinson, H., O'Gorman, J., & Shtum, D. (2016). Prospective memory and frontal lobe function. *Aging, Neuropsychology and Cognition*, 23 (2), 171-183.

Okuda, J., Fuji, T., Yamadori, A., Kawashima, R., Tsukiura, T., Fukatsu, R., Suzuki, K., Ito, M., & Fukuda, H. (1998). Participation of the frontal cortices in prospective memory: evidence from a PET study in humans. *Neuroscience letters*, 253, 127-130.

Olivers, C.N.L., Smith, S., Matthews, P., & Humphreys G.W. (2005). Prioritizing new over old: an fMRI study of the preview search task. *Human Brain Mapping*, 24 (1), 69-78.

Owen, A. (1997). The functional organization of working memory processes within the human lateral prefrontal cortex: the contribution of functional neuroimaging. *European Journal of Neuroscience*, 9, 1329-1339.

Owen, A. M. (2005). Cognitive planning in humans: New insights from the Tower of London (TOL) task. In R. Morris & G. Ward (Eds.), *The Cognitive Psychology of Planning* (pp. 135–151). New York, NY: Psychology Press.

Parsons, T. (2015) Virtual reality for enhanced ecological validity and experimental control in the clinical affective and social neurosciences. *Frontiers in Human Neuroscience*, 9, 660.

Penn, P.R., Rose, F.D., & Johnson, D.A. (2008). The use of virtual reality in the assessment and rehabilitation of executive dysfunction. In M. Oddy & A.W. Worthington (Eds.), *The Rehabilitation of Executive Disorders A Guide to Theory and Practice*. Oxford University Press.

Pullen, E., Morris, R., Kerr, S., Bullock, P. R., & Selway, R. P. (2006). Exploration of social rule violation in patients with focal prefrontal neurosurgical lesions using a virtual reality simulation. *International Journal of Disabilities and Human Development*, 5 (2), 141-146.

Rabinowitz, A.R., & Levin, H.S. (2014). Cognitive Sequelae of Traumatic Brain Injury. *Psychiatric Clinics of North America*, 37 (1), 1-11.

RasPELLI, S., Pallavicini, F., Carelli, L., Morganti, F., Pedroli, E., Cipresso, P., Poletti, B., Corra, B., Sangalli, D., Silani, V., & Riva, G. (2012). Validating the Neuro VR-Based Virtual Version of the Multiple Errands Test: Preliminary Results. *Presence*, 21 (1), 31-42.

Reitan, R. M. (1992). *Trail Making Test: Manual for administration and scoring*: Reitan Neuropsychology Laboratory.

Renison, B., Ponsford, J., Testa, R., & Jansari, A.S. (2008). The use of virtual reality in assessment of executive function following traumatic brain injury. *Brain Impairment*, 9 (1), 47.

Renison, B., Ponsford, J., Testa, R., Richardson, B., & Brownfield, K. (2012). The ecological and construct validity of a newly developed measure of executive function: The virtual library task. *Journal of the International Neuropsychological Society*, 18, 440-450.

Rizzo, A.A., Schultheis, M.T., Kerns, K., & Mateer, C. (2004). Analysis of assets for virtual reality applications in neuropsychology. *Neuropsychological Rehabilitation*, 14, 207-239.

Robbins, T.W., James, M., Owen, A.M., Sahakian, B.J., Lawrence, A.D., McInnes, L., & Rabbitt, P.M. (1998), A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: implications for theories of executive functioning and cognitive aging. Cambridge Neuropsychological Test Automated Battery. *Journal of the International Neuropsychological Society*, 4 (5) 474-90.

Robertson I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). Oops!: Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia*, 35, 747– 758.

Rolls, E.T. (1999a). *The Brain and Emotion*. Oxford: Oxford University Press.

Rolls, E.T. (1999b). The functions of the orbitofrontal cortex. *Neurocase*, 5, 301–312.

Rolls, E.T. (2000). The orbitofrontal cortex and reward. *Cerebral Cortex*, 10 (3) 284-294.

Rolls, E. T. (2002). The functions of the orbitofrontal cortex. In D. T. Stuss & R. T. Knight (Eds.) *Principles of Frontal Lobe Function* (chap. 23, pp. 354–375). New York: Oxford University Press.

Rowe, A.D., Bullock, P.R., Polkey, C.E., & Morris, R.G. (2001). 'Theory of mind' impairments and their relationship to executive functioning following frontal lobe excisions. *Brain*, 124, 101-117.

Schulman, K.I. (1997). Disinhibition syndromes, secondary mania and bipolar disorder in old age. *Journal of Affective Disorders*, 46, 175-182.

Schultheis, M. T., & Rizzo, A. A. (2001). The application of virtual reality technology in rehabilitation. *Rehabilitation Psychology*, 46, 296-311.

Shallice, T. (1982). Specific impairments of planning. *Philosophical Transactions of the Royal Society of London B*, 298, 199-209.

Shallice, T., & Burgess, P. W. (1991). Deficits in strategy application following frontal lobe damage in man. *Brain*, 114, 727-741.

Shallice, T., Burgess, P., & Robertson, I. (1996). The domain of supervisory processes and temporal organisation of behaviour. *Philosophical Transactions of the Royal Society B. Biological Sciences*, 351, 1405-1412.

Shaw, P., Bramham, J., Lawrence, E.J., Morris, R., Baron-Cohen, S., & David, A.S. (2005). Differential effects of lesions of the amygdala and prefrontal cortex on recognising facial expressions of complex emotions. *Journal of Cognitive Neuroscience*, 17, 9, 1410-1419.

Shum, D., Valentine, M., & Cutmore, T. (1999). Performance of Individuals with Severe Long-Term Traumatic Brain Injury on Time-, Event-, and Activity-Based Prospective Memory Tasks, *Journal of Clinical and Experimental Neuropsychology*, 21 (1), 49-58.

Siddiqui, S. V., Chatterjee, U., Kumar, D., Siddiqui, A., & Goyal, N. (2008). Neuropsychology of prefrontal cortex. *Indian Journal of Psychiatry*, 50 (3), 202–208.

Simblett, S.K., Ring, H., & Bateman, A. (2016). The Dysexecutive Questionnaire Revised (DEX-R): An extended measure of everyday problems with executive functioning after acquired brain injury, *Neuropsychological Rehabilitation*, 19, 1-18.

Soar, K., Chapman, E., Sivakuma, N. Jansari, A. & Turner, J.J.D. (2016). Investigating the effects of caffeine on executive functions using traditional Stroop and a new ecologically-valid virtual reality task, the Jansari assessment of Executive Functions (JEF[©]). *Appetite*, 105, 156-163.

Streiner, D.L., & Norman, G.R. (2008). *Health Measurement Scales: A practical guide to their development and use*. 4th Edition. Oxford University Press, USA.

Stuss, D.T., & Benson, D.F. (1986). *The Frontal Lobes*. Raven Press, New York.

Stuss, D., & Alexander, M.P. (2000). Executive functions and the frontal lobes: a conceptual view. *Psychological Research*, 63, 289-298.

Stuss, D. T. (2007). New approaches to frontal lobe testing. In B. L. Miller & J. L. Cummings (Eds.), *The Human Frontal Lobes. Functions and Disorders* (pp. 292-305). New York: The Guilford Press.

Stuss, D.T., & Alexander, M.P. (2007). Is there a dysexecutive syndrome? *Philosophical Transactions of the Royal Society B. Biological Sciences*, 362, 901-915.

Sweeney, S., Kersel, D., Morris, R.G., Manly, T., & Evans, J. (2010). The sensitivity of a virtual reality task to planning and prospective memory impairments: Group differences and the efficacy of periodic alerts on performance. *Neuropsychological Rehabilitation*, 20 (2), 239-263.

Szczepanski, S.M., & Knight, R.T. (2014). Insights into human behavior from lesions to the prefrontal cortex. *Neuron*, 83, 1002-1018.

Toplak, M.E., West, R.F., & Stanovich, K.E. (2013). Practitioner Review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, 54, 131-143.

Tranel, D., Bechara, A., & Denburg, N.L. (2002). Asymmetric functional roles of right and left ventromedial prefrontal cortices in social conduct, decision making and emotional processing. *Cortex*, 38 (4) 589-612.

Tranel, D., Hathaway-Nepple., & Anderson, S.W. (2007). Impaired behaviour on real-world tasks following damage to the ventromedial prefrontal cortex. *Journal of Clinical and Experimental Neuropsychology*, 29 (3) 319-332.

Wechsler, D. *Wechsler Memory Scale, Third Edition*. (1997). Psychological Corporation, San Antonio, TX .

Wechsler D. *Wechsler Memory Scale, Fourth Edition*. (2009). Pearson; San Antonio, TX.

Wechsler, D. (2011). *Wechsler Abbreviated Scale of Intelligence, Second Edition* (WASI-II). San Antonio, TX: NCS Pearson.

Wechsler, D. (2011). *Manual for the Test of Premorbid Functioning*. New York: The Psychological Corporation.

Wiemels, J., Wrensch, M., & Claus, E. (2010). Epidemiology and etiology of meningioma. *Journal of Neuro-Oncology*, 99 (3), 307-314.

Wilson, B.A., Alderman, N., Burgess. P.W., Emslie, H., & Evans, J. (1996). *Behavioural Assessment of the Dysexecutive Syndrome (BADS)*. Thames Valley Test Company.

Yesavage, J., Hoblyn, J., Friedman, L., Mumenthaler, M., Schneider, B., & O'Hara, R. (2007). Should one use medications in combination with cognitive training? If so, which ones? *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 62 (1), 11-18.

Zhang, L., Abreu, B.C., Seale, G.S., Masel, B., Christiansen, C.H., & Ottenbacher, K.J. (2003). A virtual reality environment for evaluation of a daily living skill in brain injury rehabilitation: reliability and validity. *Archives of Physical Medicine and Rehabilitation*, 84 (8), 1118-24.

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta Psychiatrica Scandinavica*, 67 (6), 361-370.



Figure 1: Screen capture of the Virtual Reality office



Figure 2: Screen capture of the Virtual Reality meeting room

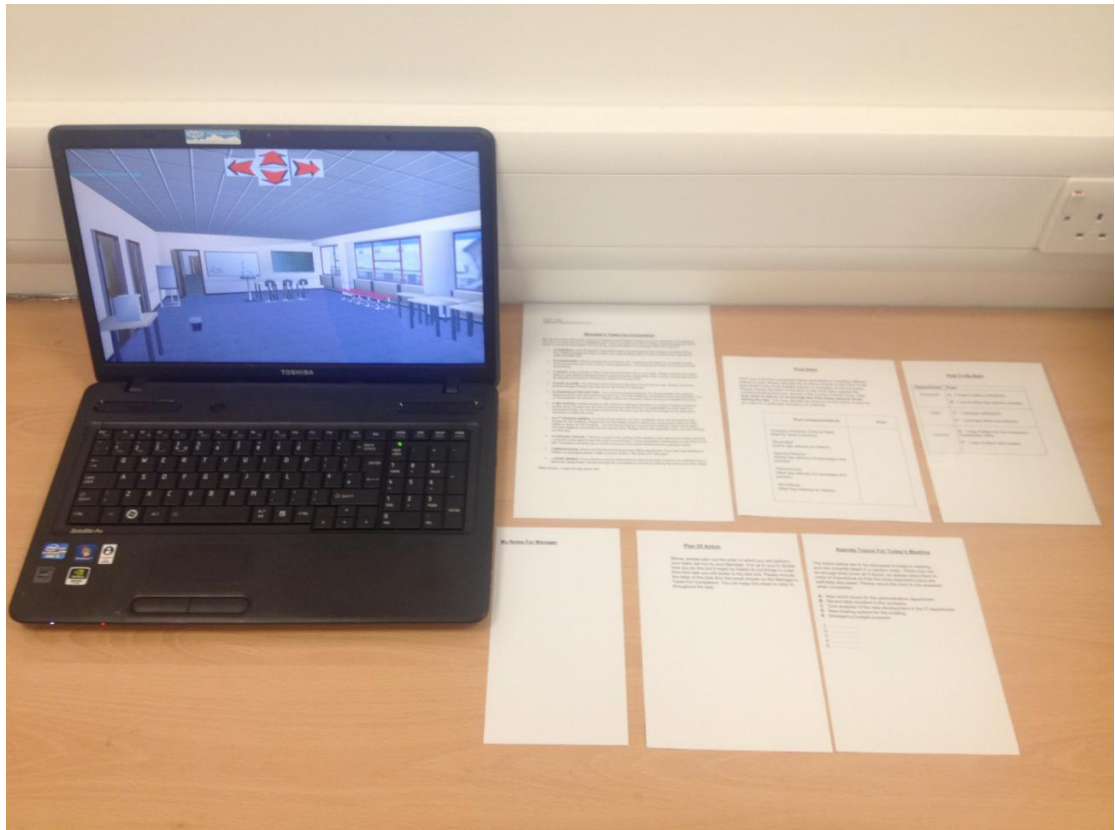


Figure 3: Setup of laptop and materials at the start of the assessment

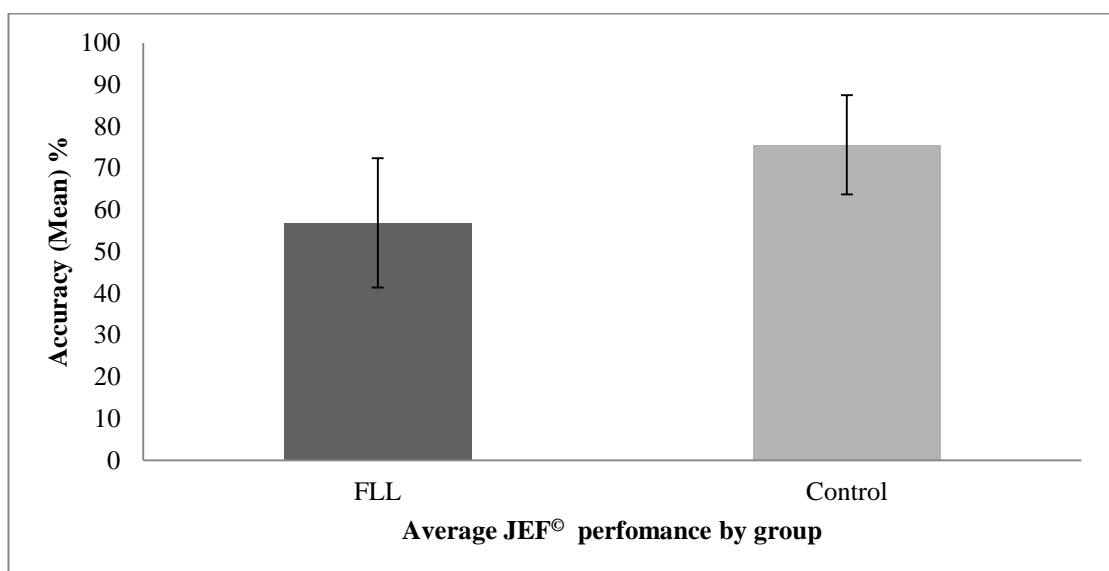


Figure 4: Average performance on the JEF[®] for the frontal lobe lesion and control groups, error bars represent the standard deviation

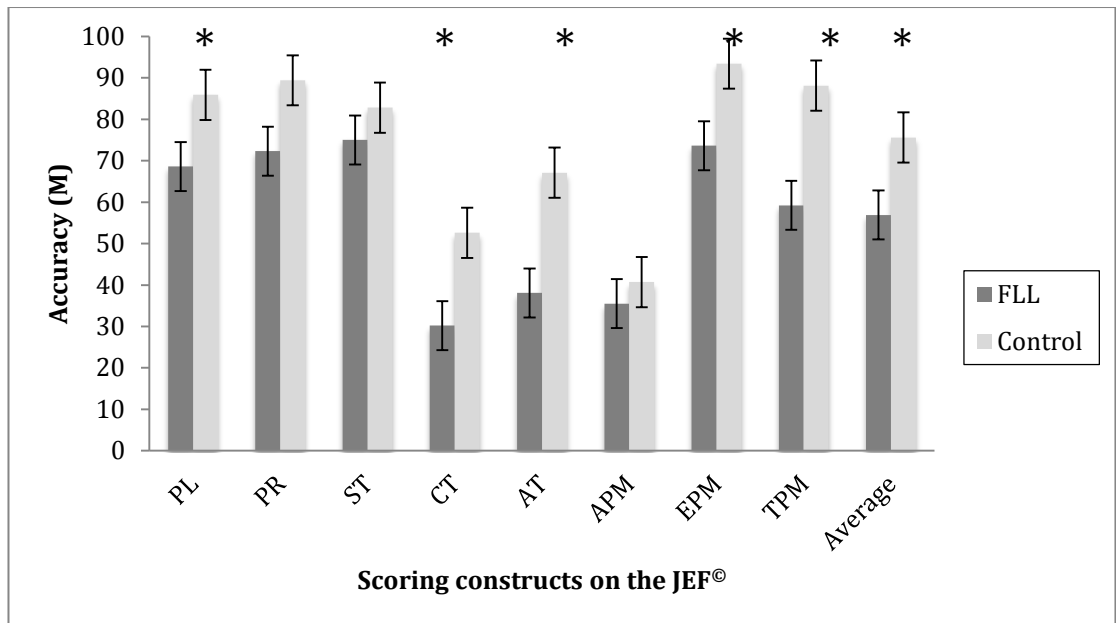


Figure 4: Performance on each construct on the JEF[®] for the frontal lobe lesion and control groups, error bars represent the standard deviation¹

¹ Construct abbreviations (PL, planning, PR, prioritisation, ST, selective-thinking, CT, creative-thinking, AT, adaptive-thinking, APM, action-based PM, EPM, event-based PM, TPM, time-based PM).

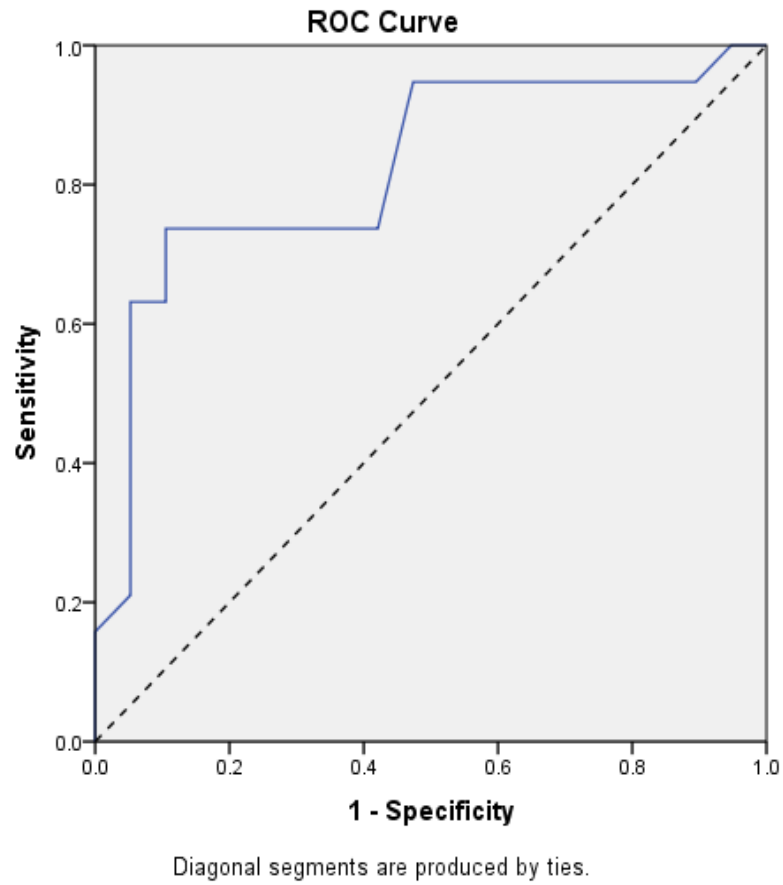


Figure 5: ROC curve for the average score on the JEF[®]. The area under the curve = 83% with a confidence interval of 0.68-0.92. Dashed line = diagonal reference line. Solid line = ROC curve